



Assessing stable validity and reliability of the Japanese version of the individualized supported employment fidelity scale: A replication

Sosei Yamaguchi¹ | Sayaka Sato¹ | Yasutaka Ojio¹ | Takuma Shiozawa¹ | Asami Matsunaga¹ | Ayano Taneda² | Utako Sawada³ | Koji Yoshida⁴ | Chiyo Fujii¹

¹Department of Community Mental Health & Law, National Institute of Mental Health, National Center of Neurology and Psychiatry, Tokyo, Japan

²Faculty of Health and Social Services, Kanagawa University of Human Services, Yokosuka, Japan

³Department of Psychiatric Nursing, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

⁴Department of Human Care and Support, Toyo University, Saitama, Japan

Correspondence

Sosei Yamaguchi, Department of Community Mental Health & Law, National Institute of Mental Health, National Center of Neurology and Psychiatry, 4-1-1 Ogawa-Higashi, Kodaira, Tokyo 187-8553, Japan.
Email: sosei.yama@ncnp.go.jp

Funding information

The present study was supported by grants to SY from the Ministry of Education, Culture, Sports, Science and Technology (Nos. 16K21661 and 20H01611) and from the Ministry of Health, Labour and Welfare (No. 19GC1006), and by grant to SS from the Japan Agency for Medical Research and Development (No. 17dk0307074h0001)

Abstract

Background: The Japanese version of the individualized Supported Employment Fidelity scale (JiSEF) was developed by modifying the 25-item Individual Placement and Support Fidelity Scale (IPS-25). While a preliminary study partly confirmed the concurrent validity with vocational outcomes, this replication study aimed to examine the stability of the concurrent validity and the inter-rater reliability of the JiSEF and to test its convergent validity with IPS-25.

Methods: Fidelity assessments were conducted in 2016 ($n = 17$), 2017 ($n = 13$), and 2018 ($n = 18$) to examine the employment rate and the fidelity scores at the agency level. We also evaluated the fidelity scores for the IPS-25 in 2018. We examined the associations between the fidelity scale scores and vocational outcomes for the concurrent validity and between the fidelity scales for convergent validity. The inter-rater reliability was examined in the 2016 and 2017 assessments.

Results: High intraclass correlation coefficients (0.93 in 2016 and 0.92 in 2017) were obtained for the inter-rater reliability. The JiSEF score in each year was associated with the agency employment rate ($r = 0.710$, $P = 0.001$ in 2016; $r = 0.722$, $P = 0.005$ in 2017; and $r = 0.665$, $P = 0.003$ in 2018). A supplementary longitudinal data analysis also confirmed the association between the JiSEF score and the employment outcomes. Additionally, the JiSEF was significantly correlated with the IPS-25 ($r = 0.760$, $P < 0.001$).

Conclusions: This study stably replicated good inter-rater reliability and concurrent validity of the JiSEF. Additionally, the convergent validity was confirmed. Further studies with large samples are needed to confirm these findings.

KEYWORDS

fidelity, mental illness, supported employment, validity, vocational outcome

A 2-3 sentence abbreviated abstract summarizing your article: The Japanese version of individualized Supported Employment Fidelity Scale (JiSEF) showed a good inter-rater reliability and concurrent validity of employment outcomes and is a promising and useful tool to implement effective employment services for people with mental illness.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2021 The Authors. *Neuropsychopharmacology Reports* published by John Wiley & Sons Australia, Ltd on behalf of the Japanese Society of Neuropsychopharmacology

1 | BACKGROUND

Supported employment for people with mental illness is an evidence-based practice. In particular, the Individual Placement and Support (IPS) model, which is well-grounded in principles, has accumulated worldwide evidence of improved vocational outcomes compared to traditional vocational services over a 30-year period across many countries including Japan.¹⁻⁸ Nevertheless, every country faces barriers to adopting supported employment such as the structure of services and relevant policies.⁹⁻¹¹ The fidelity scale is a useful tool for implementing specific practices. For supported employment, Bond and colleagues validated the 25-item Individual Placement and Support Fidelity Scale (IPS-25) and relevant studies confirmed the concurrent and predictive validity of the average agency employment rate.¹²⁻¹⁵ A study developed the Japanese version of the individualized Supported Employment Fidelity Scale (JiSEF) by partially modifying the IPS-25 to match Japan's unique mental health and labor systems.¹⁶ Preliminary validation studies reported good inter-rater reliability (intraclass correlation coefficient, ICC = 0.93) and found that the JiSEF scores were associated with the average agency employment rates ($r = 0.710$) and individual-based service intensity in Japan.

Despite the development and initial validation of the JiSEF, it is important to examine multiple aspects of the psychometric properties when developing a fidelity scale.¹⁷ The concurrent validity and inter-rater reliability of the JiSEF have each only been assessed once.^{16,18} As the IPS-25 has been rigorously validated through replication studies using cross-sectional surveys,¹²⁻¹⁴ the concurrent validity and inter-rater reliability of JiSEF need replications to confirm its stability. Furthermore, it is unclear whether the quality of implementation of supported employment assessed by the JiSEF is similar to that by IPS-25, which is the gold standard for assessing the quality of an evidence-based supported employment program. However, the convergent validity of the JiSEF has yet to be directly compared to the IPS-25. Therefore, this study aimed to reassess the stability of the concurrent validity and inter-rater reliability of the JiSEF. Additionally, it tests the convergent validity of the JiSEF compared to the IPS-25.

2 | METHODS

2.1 | Design and study setting

A cross-sectional design was employed mainly to assess the fidelity scores and unit vocational outcomes at the agency level. We conducted fidelity assessments in Japan during fiscal years 2016, 2017, and 2018. In each year, two fidelity reviewers visited each agency and conducted a fidelity review from October 1 to March 31.

2.2 | Participating agencies

In April 2016, we identified 20 agencies that potentially addressed a supported employment program in Japan either in collaboration

with the Japan IPS Association or through previous studies.^{11,16,18}

The eligibility criteria were as follows: (1) the agency offered or was interested in offering supported employment services and (2) the agency agreed to a fidelity assessment, which included interviews and reviews of service records. Each year, these 20 agencies were asked to participate in the fidelity research. Informed consent was obtained from all agencies. In 2017, the number of agencies was capped to 13 due to research grant funding limitations. The 20 agencies participated in the fidelity research at least once over a 3-year period. 11 agencies received the fidelity assessments annually for 3 years. During the study period, six agencies received the fidelity assessments twice and three agencies received it once.

2.3 | Variables and measures

2.3.1 | Agency characteristics and outcomes

Data collected from each agency included the agency's service setting, time since opening (program longevity), number of staff, employment specialists, and caseload per employment specialist, employment rate, and 3-month employment rate at the agency level. We obtained the competitive employment rate for each agency by dividing the number of users employed for one or more days by the total number of registered service users in the previous year. The 3-month employment rate was also computed using the number of service users who were employed for more than 3 months as the numerator. Competitive employment was defined as community-based work at or above the minimum wage. Employment in sheltered workplaces was not counted as a competitive employment case in this study. We also investigated the unemployment rate in each region and the population of the city where the participating agencies were located using the results of a national survey.¹⁹ and the webpage of each city.

2.3.2 | JiSEF

The JiSEF is a modified version of the IPS-25 that contains 25 items (possible scoring range 25-125) and measures each item on a 5-point behaviorally anchored scale ranging from 1 to 5. In the JiSEF assessment, reviewers obtain the necessary information for fidelity scoring through interviews with staff members such as employment specialists, case managers, and supervisors, observations by job shadowing, and access to daily service records. A higher score indicates a greater adherence to an evidence-based supported employment program. A detailed description of the JiSEF and the scale itself can be found elsewhere (Also see background and an example for item revision of the JiSEF in Online supplementary file).¹⁶

Each year, fidelity reviewers received a 1-day training that included the IPS principles, the aim of the fidelity assessment, and JiSEF evaluation and scoring methods. Specifically, the training included the interview questions of each item, data extraction from service records, and tips for scheduling job shadows and feedback.



In 2016 and 2018, a set of reviewers who worked as a service provider and as a researcher visited each agency to evaluate fidelity scores. In 2017, two reviewers who served as service providers evaluated each agency's fidelity scores. In the 2016 and 2017 assessments, the two reviewers independently evaluated the initial scores of the JiSEF to assess the inter-rater reliability. They conducted joint interviews with staff members, but they individually reviewed the service records and independently scored each item. After completion of all the JiSEF scoring, they independently entered the fidelity scores into a spreadsheet to test the inter-rater reliability. Then they compared each fidelity score and discussed the final fidelity scores of each agency. This final score was used for the concurrent validity and convergent validity. The 2018 assessment followed the same process except that the inter-rater reliability was not examined. In the 2018 assessment, the average time of the fidelity assessment was around 7 hours.

IPS-25

The IPS-25 includes 25 items (possible scoring range is 25-125). Researchers and experts in the field collaborated to develop the fidelity scale.¹² Previous studies have confirmed the predictive validity with an average agency employment rate.^{12,13} We used the IPS-25 in the 2018 assessment. One reviewer with a research background assessed fidelity to the IPS-25 for each employment program. Our research team translated the fidelity manual of the IPS-25 into Japanese²⁰ and learned fidelity scoring methods. Based on this manual, the scoring method was shared with all IPS-25 reviewers.

2.4 | Data analysis

Pearson's correlation coefficients were computed to test the associations between the fidelity scores and outcomes for the concurrent validity and between the two fidelity scales for convergent validity.¹⁶ We conducted a mixed model repeated measures for the 11 agencies that received the fidelity research in every year as a supplementary longitudinal data analysis. The model included the variables for fidelity score, time, and fidelity score and time interaction. To assess the inter-rater reliability, we computed the weighted Kappa coefficients for each item and the intraclass correlation coefficients (ICCs) for the fidelity total score. For the inter-rater reliability, each reviewer's individual score was used before their discussion on the final fidelity score of the agency. The statistical significance was set at the 5% level. All analyses were performed using Stata version 16.

3 | RESULTS

3.1 | Characteristics, outcomes, and fidelity scores

Table 1 overviews the characteristics of the participating agencies and their outcomes by year. 17 agencies participated in 2016, 13 agencies in 2017, and 18 agencies in 2018. The average agency

employment rates were 40%-45%, while the 3-month agency employment rates were around 30%. The means of the JiSEF score were 89.9 (SD = 11.8, range = 66.0-108.0) in 2016, 93.5 (SD = 13.6, range = 63.0-110.0) in 2017, and 90.1 (SD = 11.8, range = 68.0-115.0) in 2018. The mean score of the IPS-25 was 85.2 (SD = 9.3, range = 66.0-99.0).

3.2 | Inter-rater reliability

The Kappa coefficients of most items were over 0.6. Only one item "Work incentives planning" had low Kappa coefficients of 0.30 in 2016 and 0.42 in 2017. The ICC was 0.93 (95%CI, 0.74-0.98) in 2016 and 0.92 (95%CI, 0.75-0.97) in 2017 (Table 2).

3.3 | Concurrent validity and convergent validity

For all years, the JiSEF score was significantly and positively correlated with both the agency employment rate ($r = 0.710$, $P = 0.001$ in 2016; $r = 0.722$, $P = 0.005$ in 2017; and $r = 0.665$, $P = 0.003$ in 2018) and the 3-month agency employment rate ($r = 0.679$, $P = 0.003$ in 2016; $r = 0.730$, $P = 0.005$ in 2017; and $r = 0.715$, $P = 0.001$ in 2018) (Figure 1). The supplementary mixed model analysis found the significant relationship of the JiSEF score with employment rate ($B = 1.3$, 95%CI = 0.4-1.9, $P = 0.003$) and 3-month employment rate ($B = 0.9$, 95% CI = 0.2-1.6, $P = 0.009$), while the time and interaction variable were not significantly associated with the outcomes (Tables S1 and S2). In terms of the convergent validity, the JiSEF was significantly correlated with the IPS-25 in 2018 ($r = 0.760$, $P < 0.001$). The analysis also found a significant positive correlation between the IPS-25 scores and outcomes (agency employment rate, $r = 0.489$, $P = 0.039$; 3-month agency employment rate, $r = 0.541$, $P = 0.021$).

4 | DISCUSSION

We found a good inter-rater reliability, concurrent validity, and convergent validity of the JiSEF in this replication study. With regard to the inter-rater reliability, the ICC value of the JiSEF was high (>0.90) and similar to those in previous studies (ICC = 0.93).¹⁸ Based on the ICC, interpretation criteria indicated a good reliability above 0.90.²¹ The JiSEF consistently showed a good reliability when reviewers were trained. The item-level results showed a moderate to good inter-rater reliability based on a minimum acceptable Kappa value of 0.6.²² Only one item, "Work Incentive Plan," had low Kappa coefficients (0.30 and 0.42). The Japanese social security system provides a wide variety of social benefits at the national, local, and private levels for people with disabilities. Because fidelity reviewers may struggle to assess the service quality of work incentives in Japan's complex systems, further revisions may be needed for this item.

The concurrent validity of the JiSEF was confirmed using both the correlation analysis. Given the criteria of the level of correlation

TABLE 1 Characteristics, outcomes, and fidelity score of agencies in 2016-2018

| | | 2016 (n = 17) | 2017 (n = 13) | 2018 (n = 18) |
|--|-----------|-----------------------|-----------------------|-----------------------|
| Characteristics | | | | |
| Service setting | | | | |
| Psychiatric hospital/clinics | n (%) | 5 (29.4) | 2 (15.4) | 4 (22.2) |
| Community social service agency | n (%) | 12 (70.6) | 11 (84.6) | 14 (77.8) |
| Program longevity (months) | Mean (SD) | 50.5 (27.3) | 56.0 (31.3) | 71.2 (30.0) |
| Number of staff members | Mean (SD) | 7.5 (3.6) | 6.6 (3.2) | 7.2 (3.4) |
| Number of employment specialists ^a | Mean (SD) | 2.4 (1.6) | 3.0 (1.6) | 2.5 (1.4) |
| Number of caseloads per employment specialist | Mean (SD) | 24.1 (16.7) | 19.0 (15.2) | 25.0 (33.5) |
| Area unemployment rate | Mean (SD) | 3.2 (0.4) | 2.8 (0.5) | 2.6 (0.3) |
| City Population | Mean (SD) | 694,487.9 (640,500.7) | 782,421.6 (650,583.5) | 791,939.7 (692,570.4) |
| Outcomes | | | | |
| Average agency employment rate (%) | Mean (SD) | 41.0 (26.2) | 42.9 (13.1) | 45.3 (17.3) |
| Average agency employment rate: 3-months or longer (%) | Mean (SD) | 32.8 (23.3) | 27.4 (10.8) | 31.6 (19.0) |
| Fidelity scores | | | | |
| JiSEF | Mean (SD) | 89.9 (11.8) | 93.5 (13.6) | 90.1 (11.8) |
| IPS-25 | Mean (SD) | | | 85.2 (9.3) |

IPS-25, Individual Placement and Support fidelity Scale – 25 items version; JiSEF, Japanese version of individualized Supported Employment Fidelity scale.

^aFull-time equivalent.

coefficient (moderate, at least $r > 0.4$; strong, at least $r > 0.7$),^{23,24} the correlation coefficients between the JiSEF and both the average agency employment rate ($r > 0.65$) and 3-month agency employment rate ($r > 0.65$) suggest a good concurrent validity. A previous study also reported similar correlation coefficients for the average agency employment rate ($r = 0.70$) and 3-month agency employment rate ($r > 0.58$).¹⁶ The consistent results from the fidelity assessments in four different years indicate that the JiSEF's concurrent validity and applicability are stable in the Japanese mental health and labor systems. In addition, supplementary mixed model analysis for longitudinal data suggested that high-fidelity agencies maintained their higher fidelity scores and high employment rates instead of an association between the changes in the fidelity score and employment outcomes over time.

The correlation coefficients between the fidelity score and employment outcomes were higher than the IPS fidelity studies in the United States.^{12,13} There are two potential interpretations. First, the sample in this study included agencies interested in the supported employment program. However, some agencies may not actually address the program under the Japanese system, which is not suited for personalized services and integration between vocational services and mental health services.¹¹ In addition, such agencies may have a low fidelity and a low agency employment rate. This sample diversity may lead to a clear association between fidelity scores and outcomes. Another simple interpretation is that the correlation coefficient is overestimated due to the small

sample size of this study.²⁵ Therefore, a replication study with a larger sample may result in a lower correlation coefficient compared to this study.

The JiSEF was significantly correlated with the IPS-25. This result appears to support the convergent validity of the JiSEF. Since the JiSEF was developed from IPS-25, the high correlation between the scales may be unsurprising. However, a small correlation may suggest that JiSEF and IPS-25 differ substantially. In other words, the result suggests that supported employment service quality measured by the JiSEF is moderately consistent with that measured by IPS-25. Specifically, high JiSEF-fidelity agencies may, at least partially, replicate the good structural service quality of the IPS model of supported employment program, which is well-grounded in principles.

4.1 | Limitations

This study has some limitations. Despite repeated fidelity research, the sample size was small. However, most Japanese-supported employment agencies participated in this study. Due to the sample size, factor analysis and other complex longitudinal analysis for predictive validity were impossible. In addition, we experimentally used the IPS-25, in which one reviewer assessed each agency's fidelity score using a published manual instead of receiving formal training. The score reliability may be insufficient. Hence, this study may not

**TABLE 2** Inter-rater reliability of JiSEF

| | | Year 2016 | Year 2017 |
|-----|--|-----------|-----------|
| | | (n = 17) | (n = 13) |
| | | Kappa/ICC | Kappa/ICC |
| #01 | Caseload size | 0.93 | 0.88 |
| #02 | Employment services staff | 0.94 | 0.92 |
| #03 | Vocational generalists | 0.79 | 0.72 |
| #04 | Integration of rehabilitation with mental health treatment team assignment | 1.00 | 1.00 |
| #05 | Integration of rehabilitation with mental health treatment thru frequent team member contact | 0.94 | 1.00 |
| #06 | Collaboration between employment specialists and vocational rehabilitation counselors | 0.68 | 0.75 |
| #07 | Vocational unit | 0.64 | 0.69 |
| #08 | Role of employment supervisor | 0.82 | 0.66 |
| #09 | Zero exclusion criteria | 0.79 | 0.78 |
| #10 | Agency focus on competitive employment | 0.80 | 0.70 |
| #11 | Executive team support for supported employment | 0.68 | 0.67 |
| #12 | Work incentives planning | 0.30 | 0.42 |
| #13 | Disclosure | 1.00 | 0.63 |
| #14 | Ongoing, work-based vocational assessment | 0.74 | 0.70 |
| #15 | Rapid job search for competitive job | 0.80 | 0.84 |
| #16 | Individualized job search | 0.90 | 0.87 |
| #17 | Job development—Frequent employer contact | 0.87 | 0.71 |
| #18 | Job development—Quality of employer contact | 0.76 | 0.61 |
| #19 | Diversity of job type | 0.90 | 0.85 |
| #20 | Diversity of employers | 0.95 | 0.78 |
| #21 | Competitive jobs | 1.00 | 0.96 |
| #22 | Individualized follow-along supports | 1.00 | 0.70 |
| #23 | Time-unlimited follow-along supports | 0.86 | 0.70 |
| #24 | Community-based services | 0.89 | 0.95 |
| #25 | Assertive engagement and outreach by integrated treatment team | 0.79 | 0.66 |
| | Total score | 0.93 | 0.92 |

Note: ICC, Intraclass correlation coefficient; JiSEF, Japanese version of individualized Supported Employment Fidelity scale.

have validated the IPS-25 scores. Future research will need to address these issues.

5 | DATA DEPOSITORY

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to ethical restrictions. We did not explain to the participating agencies about the public access to the data in the informed consent process. When we receive a reasonable request, the data will be made available after approval by the ethics committee of the National Center of Neurology and Psychiatry.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

SY and SS conceived and designed the study. SY, SS, YO, TS, AM, AT, US, KY data acquisition were responsible for data acquisition. SY analyzed data. KY and CF provided administrative and technical supports. All the authors contributed to the interpretation of the results, writing the manuscript and approved the final version of the manuscript.

APPROVAL OF THE RESEARCH PROTOCOL BY AN INSTITUTIONAL REVIEWER BOARD

This study was approved by the Research Ethics Committee at the National Centre of Neurology and Psychiatry (A2016-055).

INFORMED CONSENT

Informed consent was obtained from all agencies.

Animal Studies: n/a.

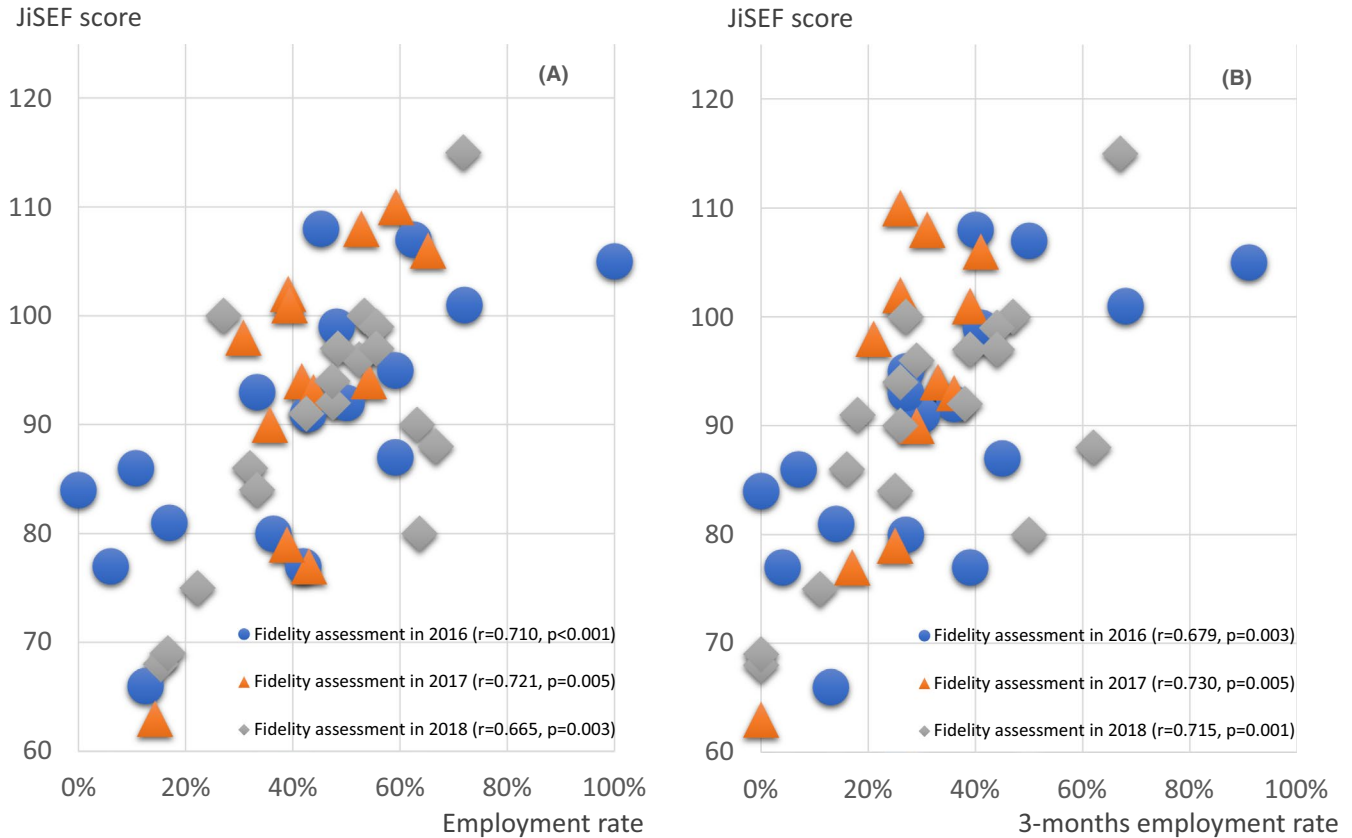


FIGURE 1 Pearson's correlation coefficient analysis for concurrent validity. (A) Correlation between JiSEF score and employment rate at the agency level; (B) correlation between JiSEF score and 3-month employment rate at the agency level. Employment rate, Average agency employment rate; 3-months employment rate, Average agency employment rate: 3-months or longer; JiSEF, Japanese version of individualized Supported Employment Fidelity scale

REGISTRY AND THE REGISTRATION NO. OF THE STUDY/TRIAL

n/a.

ORCID

Sosei Yamaguchi  <https://orcid.org/0000-0002-0579-4431>

Sayaka Sato  <https://orcid.org/0000-0002-3386-9576>

Yasutaka Ojio  <https://orcid.org/0000-0002-9660-206X>

Takuma Shiozawa  <https://orcid.org/0000-0002-9307-576X>

REFERENCES

- Modini M, Tan L, Brinchmann B, Wang M-J, Killackey E, Glozier N, et al. Supported employment for people with severe mental illness: systematic review and meta-analysis of the international evidence. *Br J Psychiatry*. 2016;209:14–22.
- Suijkerbuijk YB, Schaafsma FG, van Mechelen JC, Ojajärvi A, Corbière M, Anema JR. Interventions for obtaining and maintaining employment in adults with severe mental illness, a network meta-analysis. *Cochrane Database Syst Rev*. 2017;9:CD011867.
- Metcalfe JD, Drake RE, Bond GR. Economic, labor, and regulatory moderators of the effect of individual placement and support among people with severe mental illness: a systematic review and meta-analysis. *Schizophr Bull*. 2018;44:22–31.
- Brinchmann B, Widding-Havneraas T, Modini M, Rinaldi M, Moe CF, McDaid D, et al. A meta-regression of the impact of policy on the efficacy of individual placement and support. *Acta Psychiatr Scand*. 2020;141:206–20.
- Frederick DE, VanderWeele TJ. Supported employment: meta-analysis and review of randomized controlled trials of individual placement and support. *PLoS One*. 2019;14:e0212208.
- Richter D, Hoffmann H. Effectiveness of supported employment in non-trial routine implementation: systematic review and meta-analysis. *Soc Psychiatry Psychiatr Epidemiol*. 2019;54:525–31.
- Yamaguchi S, Sato S, Horio N, Yoshida K, Shimodaira M, Taneda A, et al. Cost-effectiveness of cognitive remediation and supported employment for people with mental illness: a randomized controlled trial. *Psychol Med*. 2017;47:53–65.
- Oshima I, Sono T, Bond GR, Nishio M, Ito J. A randomized controlled trial of individual placement and support in Japan. *Psychiatr Rehabil J*. 2014;37:137–43.
- Kirsh B. Client, contextual and program elements influencing supported employment: a literature review. *Community Ment Health J*. 2016;52:809–20.
- Becker DR, Bond GR. Commentary on special issue on individual placement and support (IPS) international. *Psychiatr Rehabil J*. 2020;43:79–82.
- Hayashi T, Yamaguchi S, Sato S. Implementing the individual placement and support model of supported employment in Japan: barriers and strategies. *Psychiatr Rehabil J*. 2020;43:53–9.
- Bond GR, Peterson AE, Becker DR, Drake RE. Validation of the revised individual placement and support fidelity scale (IPS-25). *Psychiatr Serv*. 2012;63:758–63.



13. Kim SJ, Bond GR, Becker DR, Swanson SJ, Langfitt-Reese S. Predictive validity of the individual placement and support fidelity scale (IPS-25): a replication study. *J Vocat Rehabil.* 2015;43:209–16.
14. Lockett H, Waghorn G, Kydd R, Chant D. Predictive validity of evidence-based practices in supported employment: a systematic review and meta-analysis. *Ment Health Rev J.* 2016;21:261–81.
15. de Winter L, Couwenbergh C, van Weeghel J, Bergmans C, Bond GR. Fidelity and IPS: does quality of implementation predict vocational outcomes over time for organizations treating persons with severe mental illness in the netherlands? *Soc Psychiatry Psychiatr Epidemiol.* 2020;55:1607–17.
16. Sasaki N, Yamaguchi S, Shimodaira M, Sato S, Taneda A, Yoshida K, et al. Development and validation of a Japanese fidelity scale for supported employment. *Adm Policy Ment Health.* 2018;45:318–27.
17. Bond GR, Drake RE. Assessing the fidelity of evidence-based practices: history and current status of a standardized measurement methodology. *Adm Policy Ment Health.* 2020;47(6):874–84.
18. Yamaguchi S, Mizuno M, Sato S, Matsunaga A, Sasaki N, Shimodaira M, et al. Contents and intensity of services in low- and high-fidelity programs for supported employment: results of a longitudinal survey. *Psychiatr Serv.* 2020;71:472–9.
19. Statistics Bureau. Labor Force Survey (basic tabulation): Results by prefecture. Ministry of Internal Affairs and Communications. 2020. Available at: <https://www.stat.go.jp/data/roudou/pref/index.html>. Accessed 31th March
20. Becker DR, Swanson S, Bond GR, Merrens MR. Evidence-based supported employment fidelity review manual, 2nd edn. Hanover: Dartmouth Psychiatric Research Center; 2011.
21. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15:155–63.
22. McHugh ML. Interrater reliability: the kappa statistic. *Biochemia Medica.* 2012;22:276–82.
23. Mukaka MM. Statistics corner: a guide to appropriate use of correlation coefficient in medical research. *Malawi Med J.* 2012;24:69–71.
24. Akoglu H. User's guide to correlation coefficients. *Turk J Emerg Med.* 2018;18:91–3.
25. Makin TR, Orban de Xivry J-J. Ten common statistical mistakes to watch out for when writing or reviewing a manuscript. *Elife.* 2019;8:e48175.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Yamaguchi S, Sato S, Ojio Y, et al. Assessing stable validity and reliability of the Japanese version of the individualized supported employment fidelity scale: A replication. *Neuropsychopharmacol Rep.* 2021;41:248–254. <https://doi.org/10.1002/npr2.12172>